

MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU OF STANDARDS -1963 - A

# PLANNING UNDER INCOMPLETE INFORMATION AND THE RATCHET EFFECT

ру

Xavier Freixas, Roger Guesnerie and Jean Tirole

Technical Report No. 408

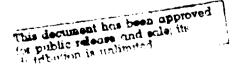
June 1983

A REPORT OF THE
CENTER FOR RESEARCH ON ORGANIZATIONAL EFFICIENCY
STANFORD UNIVERSITY
Contract ONR-N0014-79-C-0685, United States Office of Naval Research



THE ECONOMICS SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
Fourth Floor, Encina Hall
Stanford University
Stanford, California
94305



# PLANNING UNDER INCOMPLETE INFORMATION AND THE RATCHET EFFECT\*

by

Xavier Freixas\*\*, Roger Guesnerie\*\*\* and Jean Tirole\*\*\*

## 1. Introduction

metric information. The firm in general has more information about its productive possibilities than the Central Planner (from now on for short CP). This justifies the use of incentives schemes in which the CP does not directly fix activity levels. Casual observation of such existing schemes suggests that the time dimension is crucial for their working. Typically the CP revises the incentive scheme over time to take into account the information provided by the firm's performance. Managers of centrally planned economies as well as economists have long recognized that this revision induces firms to underproduce in order to avoid demanding schemes in the future. This is the ratchet effect.

Although the study of incentives schemes in centrally plannaed economies has become popular in the late seventies, few researches have focused on the ratchet effect. Various authors have investigated how schemes may induce the firm to reveal its information before producing (Fan [1975], Bonin [1976], Gindin [1970]) or the effect of a given

N

Avell and/or

<sup>\*</sup>This research was supported in part by the Office of Naval Research Contract ONR-N00014-79-C-0685 at the Institute for Mathematical Studies in the Social Sciences at Stanford University.

<sup>\*\*</sup>Universite des Sciences Sociales, Toulouse, France.

<sup>\*\*\*</sup>CERAS and EHESS, Paris, France.

<sup>\*\*\*\*</sup>CERAS, Paris, France.

incentive scheme on the firm's behavior (see Weitzman [1976] on the new Soviet Incentive Scheme). Also Weitzman [1980] gave examples of the ratchet effect assuming a given (not necessarily optimal) intertemporal scheme. Holmstrom [1979] showed that the revision procedures in the new Soviet Incentive Scheme in general dominate fixed target schemes. The two latter contributions rely on the assumption that the CP commits himself in advance to a given scheme. Although the exploration of the phenomenon under this latter assumption is useful, it must be stressed that most often, in practice, the CP does not commit himself to a revision procedure; this latter fact seems to be crucial for understanding qualitative features of observed ratchet effects.

It is the purpose of the present paper to study the ratchet effect under the non-commitment assumption, which we believe to be more realistic. The inadequacy of conceptual tools made it difficult in the past to consider such an assumption. Now basic research on dynamic games with incomplete information recently resulted in new ideas which look particularly relevant for a theoretical analysis of the problem.

The dynamic theory of the relationship CP-firm that we attempt here, relies on these modern game-theoretic developments. Although our model is highly stylized, it is designed to incorporate the central positive and normative aspects of the problem, as they have been more or less explicitly described in previous theoretical literature.

The content of the paper can be summarized as follows: Section 2 introduces the model. The CP is assumed to have incomplete information about the productivity of the firm (in most of our paper, we shall

suppose that this productivity can take only two values). We explain why the CP does not decentralize the "first best" optimum by having the firm face the social value of its output. Hence, in a second best framework we consider two classes of schemes: the linear and the totally non-linear ones. As a starting point, we study the influence of incomplete information on the optimal incentive scheme in a static framework (Section 3). We show that the marginal price of output in the reward scheme is lower than its shadow price and we explain why. Section 4 introduces dynamics and the possibility for the CP to revise the incentive scheme over time. To simplify we consider a two-period model where the essence of the phenomenon can be captured while letting the argument be reasonably simple. To be consistent with casual observation, we assume that the CP does not commit himself to a intertemporal incentive scheme at the start. Reasons why commitment may not be possible are discussed in Section 7. The sequential choice of reward schemes by the CP and outputs by the firm is most naturally modelled as a game between the two partners. We show that this game has a unique welfare outcome. In the optimal incentive scheme, the firm may hide its information or reveal it. Even in the latter case the ratchet effect exists in the sense that the CP may choose a scheme which is suboptimal from a static point of view in order to induce revelation. We also show that we should expect the marginal price of output in the reward scheme to exceed its optimal static marginal price. These results are derived in Section 5.

Most of our analysis is concerned with (optimal) linear and non-

linear schemes. An important part of the recent literature deals with "intermediate" schemes, the so-called old and new Soviet incentive schemes. Section 6 discusses these schemes in the light of our model.

It is likely that the relevance of this work goes far beyond Central Planning since any type of long run relationship between two agents under asymmetric information, delegation within a firm, taxation, etc., is not one-shot. As the principal then learns necessarily over time, some form of ratchet effect necessarily obtains. Section 8 mentions these applications and stresses analogies. But much work remains to be done to formalize them.

### The Model

## 2.1 The Set Up

a) <u>Technology</u>: A firm produces one good whose "social value" per unit p is exogenously given. The production parameter y is one-dimensional. It can be interpreted either as the production level or as excess production above a minimum level which can be enforced through existing direct controls. The cost to the firm!/ of producing y (excess) units of the good is a function  $\psi(y,\theta)$  of y and of a productivity parameter  $\theta$ ;  $\theta$  stands for the ability of workers and managers, true capacity of the firm, etc., and can only be observed by the firm. The CP has some probability distribution  $f(\theta)$  about it, which is common knowledge. In most of the paper, the probability distribution will be confined to two values  $\theta$  and  $\theta(\theta < \theta)$  with respective probabilities  $\nu$  and  $(1 - \nu)$ . The cost function will then be

written  $\psi(y)$  or  $\overline{\psi}(y)$ .

The production cost can be thought of as embodying the managers and workers' effort and cannot be observed by the CP. We shall make the following assumptions on  $\psi\colon$  the cost increases with and is convex in output:  $\psi_y<0$ ,  $\psi_{yy}>0$  and decreases with productivity:  $\psi_\theta<0$ . Furthermore the marginal cost of production decreases with productivity:  $\psi_{v\theta}<0$ .

b) Incentive Scheme: The CP rewards the firm with an incentive scheme R(y), and derives a utility py from output y.

Standard first best analysis would suggest to reward the firm according to R(y) = py; this conclusion will be incorrect here where it is assumed that there is a social cost of  $\lambda$  to transferring \$1 to the firm. There are two possible polar interpretation of this shadow price  $\lambda$ .

First, \$1 given to the firm is worth, in terms of social welfare  $\$(1-\lambda);$  the multiplier  $\lambda$  then accounts for distortions associated with the firm's distribution of its revenue. This interpretation looks particularly pertinent to the context of Central Planning; such an organization, particularly in its soviet-type version, stems from a distrust of market mechanisms, both for their distributional effects on individual incomes but also for their allocational effects on investments. The distribution of proceeds py, would have undesirable features when coming in addition to existing money flows in an otherwise planned economy; the surplus if distributed may generate uncontrolled inequalities or if it is invested, lead to relative capital accumulation

considered as unappropriate. For a more basic reflection on these phenomena, a general equilibrium analysis would be needed; the multiplier should then appear as the outcome of a second best optimization under basic informational and derived institutional constraints (absence of capital markets, etc.).

A second interpretation relates  $\lambda$  to the social cost of raising funds. This is a standard assumption, precise justifications of which have been given when, for example, funds are raised through distortionary taxation. Justifications of a similar kind could be given in the context of Central Planning.

In formulating the problem just in the following, we will stick, for the sake of convenience, to the second interpretation, although the reader will convince himself that the first story or any mixture of the first and the second stories can lead to the same formulation.

We formalize the firm's objective function  $\pi^{\rm F}$  as the difference between its revenue and its cost:

(2.1) 
$$\pi^{F}(y,\theta) = R(y) - \psi(y,\theta)$$

It is convenient to split the CP's objective function into two parts. First, the CP is concerned with the "rest of the world" welfare, i.e., with the difference between the social value of output and the social cost of transferring R(y) to the firm. This first part will be referred to as the CP's "social concern" utility function and denoted  $\pi^{CP}$ . Sticking to the interpretation of  $\lambda$  as a shadow cost of raising \$1:

(2.2) 
$$\pi^{CP}(y) = py - (1 + \lambda)R(y)$$

The CP also takes into account the firm's welfare  $\pi^F$ . Thus aggregate welfare W can be written:

(2.3) 
$$W(y,\theta) = \pi^{F} + \pi^{CP} = py - \psi(y,\theta) - \lambda R(y)$$

It is useful to represent indifference curves for  $\pi^F$  and  $\pi^{CP}$  in the revenue-output space. This is done in Figures 1 and 2. We shall assume that the CP is risk neutral. We also assume that the firm can guarantee itself some minimum utility, for example, as suggested above by producing at a minimum acceptable level. Without loss of generality we shall assume that the minimum utility level is associated with a zero production and is equal to 0. Thus the firm maximizes  $\pi^F(y,\theta)$  subject to the individual rationality constraint  $\pi^F(y,\theta) > 0$ . If  $\max_{y \in Y} \pi^F(y,\theta) < 0$  the firm's output is zero.

Let s denote the ratio  $p/(1 + \lambda)$ . Note that s is the slope of the CP's social concern indifference curves. For reasons that will become clear shortly, s will be called the shadow price of cutput.

An important part of this paper will be concerned with the class of linear schemes, i.e., schemes of the following type: R(y) = a + by. A linear scheme can also be written as:  $R(y) = b(y - y_0)$  where  $y_0 = -(a/b)$  can be interpreted as a quota, and b as a bonus per unit of output above the quota. 3/

### 2.2 Static Incentive Scheme Under Complete Information

In order to motivate the following, we first consider the simple case of complete information about  $\theta$ . Let us derive the optimal linear scheme (which in this case is also the optimal non-linear scheme). Let

 $y(b,\theta)$  denote the output that maximizes  $\pi^F(y,\theta)$ . For any given bonus b, it is desirable from the viewpoint of aggregate welfare to choose the lump-sum transfer a so as to make the firm's individual rationality constraint binding: lowering it by  $\delta a > 0$ , other things being equal, does not affect the firm's production decision and increases aggregate welfare W by  $\delta W = \lambda \delta a$ . Hence

 $a = \psi(y(b,\theta),\theta) - by(b,\theta)$ . What bonus will the CP choose? b must be such that the (marginal) social price of output net of transfer costs equals the marginal value of output for the firm (b). The former is nothing but the difference between the social value of one unit of output p and the shadow cost of transferring \$b,  $\lambda b$ . Thus

$$(2.4) p - \lambda b = b$$

or

(2.5) 
$$b = \frac{p}{1 + \lambda} = s$$

The formal proof of this latter point is as simple as the heuristics: given that we have shown that the individual rationality constraint of the firm has to be binding at the optimum, the CP's problem reduces to the maximization of  $\{sy-R\}$  over the indifference curve of the firm  $R-\psi(y,\theta)=0$ , i.e., to the unconstrained maximization of  $sy-\psi(y,\theta)$ . The optimum is uniquely determined by  $\psi_y=s$  and is decentralized through a linear bonus function with b=s.

Figure 1:  $\theta < \overline{\theta}$ 

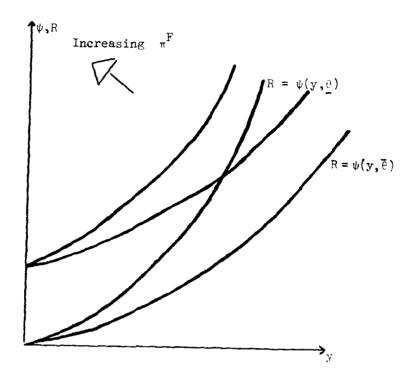
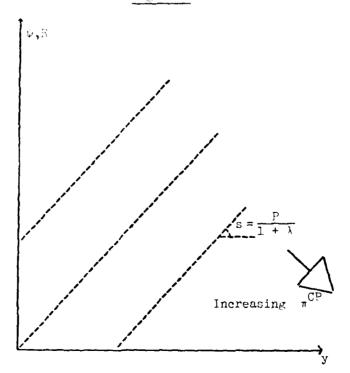


Figure 2



It follows:

<u>Proposition 1</u>: The optimal bonus under complete information is equal to the shadow price of output s.

## 3. Static Incentive Schemes Under Incomplete Information

We first analyze the optimal linear scheme when the relationship between the firms and the CP is short-run and there are two types of firms. We then test the robustness of our main conclusion to the class of eligible schemes (by allowing non-linear ones) and to the distribution of firms.

a) Linear Scheme With Two Types of Firms: We assume that the CP can choose only in the class of linear schemes and that the firm has two potential productivities  $\theta$  and  $\overline{\theta}$  with probability  $\nu$  and  $(1-\nu)$ . We shall also assume that the proportion of low productivity firms is not "too low" relative to the productivity differential, so that the CP never wants to induce such a firm not to operate (otherwise the problem is trivial). Let  $\underline{y}(b)$  and  $\overline{y}(b)$  be the firm's utility-maximizing outputs of both types of firms when the bonus is b. Note that  $\overline{y}(b) > y(b)$ . The CP maximizes:

$$(3.1) \qquad \mathbf{W} = \mathbf{v} \left[ \mathbf{p} \underline{\mathbf{y}}(\mathbf{b}) - \psi(\underline{\mathbf{y}}(\mathbf{b})) - \lambda(\mathbf{a} + \mathbf{b} \underline{\mathbf{y}}(\mathbf{b})) \right]$$

$$+ (1 - \mathbf{v}) \left[ \mathbf{p} \overline{\mathbf{y}}(\mathbf{b}) - \overline{\psi}' \overline{\mathbf{v}}(\mathbf{b}) \right] - \lambda(\mathbf{a} + \mathbf{b} \overline{\mathbf{y}}(\mathbf{b})) \right]$$

Clearly, the lump-sum transfer must make the low-productivity firm's individual rationality constraint binding (the argument of Section 1 is transposed here):

(3.2) 
$$a = \psi(y(b)) - by(b)$$

Differentiating (3.1) gives:

(3.3a) 
$$\frac{dW}{db} = (1 + \lambda)[(s - b)[v \frac{dy}{db} + (1 - v) \frac{d\overline{y}}{db}] - \frac{\lambda(1 - v)}{1 + \lambda}[\overline{y}(b) - \underline{y}(b)]]$$

The first order condition obtains when:

(3.3b) 
$$s - b = \frac{\lambda(1 - \nu)}{1 + \lambda} \left[ \frac{\bar{y}(b) - \underline{y}(b)}{\nu \frac{d\underline{y}}{db}(b) + (1 - \nu) \frac{d\underline{y}}{db}(b)} \right]$$

Since output grows with the productivity and the bonus, we conclude from (3.3a,b):

<u>Proposition 2</u>: Under incomplete information about the productivity of the firm,

- (i) W is a decreasing function of b for b > s.
- (ii) the optimal bonus is smaller than the shadow price of output

We can give a more intuitive proof of this proposition as follows: Assume that for the optimal scheme b > s. Let  $B = (R, \bar{y})$ ,  $B = (R, \bar{y})$  be the choices of firms  $\bar{\theta}$  and  $\bar{\theta}$ . Now assume that the CP offers the linear scheme with slope s that leaves the low productivity firm with a zero surplus. If the firm is of type  $\bar{\theta}$ , it is indifferent to this change of scheme; since the CP reaches a better "social concern" indifference curve, ex-post aggregate welfare, in case the firm has a low productivity, increases. Assume now that the firm has type  $\bar{\theta}$ . The firm reacts to the new scheme of slope s by choosing B' = (R', y').

This change can be decomposed into two steps.

- A change along ''e indifference curve through the original point (B,R). This corresponds to a "substitution effect". For the same reason, as above, the B firm being indifferent to this move and the CP's social concern utility increasing, aggregate welfare increases.
- A change from the upper indifference curve through B to the lower indifference curve through B'. This corresponds to an "income effect" and given our separability assumption it only involves an income transfer from the firm to the society and according to (2.3) this move is beneficial from the point of view of social welfare.

Hence if the firm is of type  $\overline{\theta}$ , aggregate welfare increases when moving from a bonus scheme b > s to a bonus scheme with b = s. Hence b > s cannot be optimal. The argument conveys the following intuitive ideas: The CP is forced to leave a surplus to the high productivity firm if he wants the low productivity firm to produce. Raising the bonus above the shadow price of output creates a productive misallocation and increases the high productivity firm's surplus (see below) and thus cannot be optimal.  $\frac{b}{4}$ 

b) Varying the Probability Distribution of Firms: It will be very useful for the dynamic study to consider what happens when the proportion of high and low productivity firms changes. Let  $\bar{\pi}^F(\nu)$  denote the high productivity firm's profit when the proportion of low productivity firms is  $\nu$  and the CP chooses the optimal linear scheme  $(a(\nu), b(\nu))$  (so as to keep both types active). Remember that the profit of the low productivity firm  $\pi^F$  is equal to 0.

We can show:

### Proposition 3:

- (i) b(v) increases with v.
- (ii)  $\bar{\pi}^{F}(v)$  is continuous and increasing in v.

Proposition 3, which is proved in Appendix 1, says that the optimal bonus and the profit of high productivity firms increase with the probability of a low productivity firm. This latter fact is indeed the basic reason why the firm would like to pretend it has a productivity lower than its true one.

c) Optimal Non-Linear Scheme: We continue with the two-type case, but now with a non-linear scheme. The kind of reasoning involved here is reminiscent of the analysis on optimal taxation originating in Mirrlees [1971] and in particular is closely related to the analysis of Guesnerie-Seade [1982] which is concerned with a finite number of types.5/

The CP chooses two pairs of output-reward  $(\underline{y},\underline{R})$  for firm  $\underline{\theta}$  and  $(\bar{y},\overline{R})$  for firm  $\overline{\theta}$ . These must satisfy two conditions: Self-selection (no type wants to choose the other type's pair) and individual rationality (both firms are free not to operate). As is usual, the only binding self-selection constraint corresponds to the high productivity firm not pretending it has a low productivity. Thus:

$$(3.4) \bar{R} - \bar{\psi}(\bar{y}) = \bar{R} - \bar{\psi}(\underline{y})$$

Furthermore, the low productivity firm must have no surplus.

$$R = \underline{\psi}(\underline{y})$$

Otherwise the CP could transfer money to himself by reducing  $\frac{R}{r}$ , and this would only relax the self-selection constraint.

Note that we reduced the CP's problem to the choice of two outputs,  $\underline{y}$  determines  $\underline{R}$  from (3.5) and (3.4) then gives  $\overline{R}$  given  $\overline{y}$ . Thus the CP maximizes:

(3.6) 
$$\max_{\{\underline{y}, \overline{y}\}} \{ v[\underline{p}\underline{y} - \underline{\psi}(\underline{y})(1 + \lambda)] + (1 - v)[\underline{p}\overline{y} - \overline{\psi}(\overline{y}) - \lambda(\underline{\psi}(\underline{y}) + \overline{\psi}(\overline{y}) - \overline{\psi}(\underline{y})] \}$$

The first order conditions for this problem are:

$$(3.7) \qquad \qquad \bar{\psi}'(\bar{y}) = s$$

(3.8) 
$$s - \underline{\psi}'(\underline{y}) = \frac{\lambda(1-\nu)}{(1+\lambda)\nu} [\underline{\psi}'(\underline{y}) - \overline{\psi}'(\underline{y})]$$

(3.7) says that the high productivity firm locally faces the shadow price of output s. This is the familiar "no-distortion-at-the-top" result. (3.8) is somewhat similar to (3.3a,b). Since the marginal cost of production decreases with productivity, (3.8) implies that the marginal cost for the low productivity firm is lower than the shadow price of output s. In this sense, Proposition 2 is robust when one considers non-linear schemes.

# 4. Long-Run Relationship and the Ratchet Effect: Existence of Equilibrium

a) The Dynamic Model: We now want to allow for the possibility that the CP learns over time about the firm's constant productivity.

Our framework is reminiscent of that employed by Fudenberg and Tirole [1983] in their study of sequential bargaining under incomplete information. There are two types of firms,  $\theta$  and  $\theta(\theta < \theta)$  with respective probabilities  $v_1$  and  $(1 - v_1)$ . There are two periods, t = 1,2. The social price of output p remains constant over time. We first assume that the CP chooses schemes in the linear class. The timing is the following: At the start of the first period the CP chooses a scheme  $(a_1,b_1)$ ; the firm reacts to this scheme by producing  $y_1$  at cost  $\psi(y_1,\theta)$  and is rewarded  $(a_1 + b_1y_1)$ . Then at the start of the second period the CP chooses a new scheme  $(a_2,b_2)$ , and the firm chooses a second period output  $y_2$ . Let  $\delta$  be the common discount factor.  $\frac{\delta}{2}$  With straightforward notation the discounted profit and aggregate social welfare are:

$$(4.1) \quad \pi^{F} = \pi_{1}^{F} + \delta \pi_{2}^{F} = [a_{1} + b_{1}y_{1} - \psi(y_{1}, \theta)] + \delta[a_{2} + b_{2}y_{2} - \psi(y_{2}, \theta)]$$

$$\begin{aligned} (4.2) \quad & w = w_1 + \delta w_2 = (\pi_1^F + \delta \pi_2^F) + (\pi_1^{CP} + \delta \pi_2^{CP}) \\ & = [py_1 - \psi(y_1, \theta) - \lambda(a_1 + b_1y_1)] + \delta[py_2 - \psi(y_2, \theta) - \lambda(a_2 + b_2y_2)] \end{aligned}$$

The situation depicted above is a dynamic game between the firm and the CP. The CP's strategy is a sequence of schemes  $(a_1,b_1)$  and  $(a_2,b_2)(a_1,b_1,y_1)$  where the second period scheme is a function of his information at that date, i.e., that the firm has chosen to produce  $y_1$  under the incentive scheme  $(a_1,b_1)^{-1}$  (we insist on the fact that the CP does not commit himself in advance to a given revision process). The firm's strategy is a sequence of outputs  $y_1(a_1,b_1,\theta)$  and  $y_2(a_1,b_1,a_2,b_2,\theta)$ , where the

output at a given period is a function of the firm's information at that date (clearly,  $y_2$  will depend only on  $(a_2,b_2)$  and  $\theta$  in equilibrium). Mixed strategies are allowed. Po introduce the equilibrium notion, we must also describe the CP's beliefs about the productivity of the firm at the beginning of the second period. Let  $v_2(a_1,b_1,y_1)$  denote the CP's posterior probability that the firm has a low productivity. The equilibrium notion - Perfect Bayesian, or Sequential Equilibrium in the terminology of Kreps-Wilson [1982] to which we refer for a careful analysis - requires that strategies and beliefs be consistent; i.e., that the strategies be optimal given the beliefs and that the beliefs be derived from strategies using Bayes rule.

A <u>Perfect Bayesian Equilibrium</u> is a set of (possibly mixed) strategies  $\{(a_1,b_1), y_1, (a_2,b_2), y_2\}$  and of beliefs  $\{v_2\}$  satisfying (we shall often omit the arguments of strategies and beliefs for clarity):

- P1)  $\forall \theta, y_2$  maximizes  $\pi_2^F$
- P2)  $(a_2,b_2)$  maximizes the expectation of  $W_2$ , given the CP's beliefs  $v_2(a_1,b_1,y_1)$  and the firm's second period strategy  $y_2$ .
- P3)  $\forall \theta, y_{\eta}$  maximizes  $\pi^{F}$  given the second period strategies.
- P4) (a<sub>1</sub>,b<sub>1</sub>) maximizes the expectation of W, given the firm's and the CP's subsequent strategies.
- B)  $v_2(a_1,b_1,y_1)$  is Bayes-consistent with the prior probability  $v_1$  and the firm's first period strategy  $y_1(a_1,b_1,\theta)$ .
- (P1) to (P4) only describe the principles of Kuhn's algorithm for

finding a dynamic perfect Nash equilibrium when the probabilities of reaching the knots of the tree are exogenously given. (B) endogenizes these probabilities in conformity with Bayes rule.

### b) Resolution:

Note that a given {a<sub>1</sub>,b<sub>1</sub>} defines a game between the CP and the firm which is a subform of the initial game. Perfectness requires that equilibrium strategies in the larger initial game induce an equilibrium for this game, which we call a continuation equilibrium.

We will say that a continuation equilibrium is unique if the observed equilibrium actions are unique.

In this section, we focus on existence and uniqueness of continuation equilibria. We shall denote by  $\underline{y}(b_1)$  and  $\overline{y}(b_1)$  the outputs that maximize the <u>one-period</u> surplus of the low and high productivity firm when the bonus is  $b_1$  and  $\underline{y}_1(b_1)$  and  $\overline{y}_1(b_1)$  the <u>first-period</u> outputs when the firms take into account the effect of their choice on the second period incentive schemes; under the assumption that the firms operate, these outputs do not depend on  $a_1$ , as we shall see below.

We first notice that from (P1) and (P2) the second period actions of the firm and the CP are the same as in the one-period game with prior  $v_2$ . They have been analyzed in Section 3. Let us now consider the first period actions. We first assume that both types of firms operate (we come back to this assumption later). For the moment, the statement "for any bonus  $b_1$ " should be qualified by "assuming that  $a_1$  is high enough so that both firms operate".

We first derive two <u>necessary</u> conditions for continuation equilibrium.

$$\underline{\text{Lemma 1}}: \quad \forall \quad b_1: \quad \underline{y}_1(b_1) = \underline{y}(b_1)$$

In other words, the optimal first period action of the low productivity firms, in the dynamic context and in equilibrium, can only be its optimal one-period action.

Proof of Lemma 1: From Section 3, we know that, whatever his beliefs  $v_2$ , the CP will not allow the low productivity firm to make a strictly positive surplus in the second period. Thus this type of firm is only concerned with its first period surplus  $\underline{\pi}_1^F$ , and chooses  $\underline{y}_1(b_1)$  so as to maximize it. Q.E.D.

$$\underline{\text{Lemma 2}} : \quad \forall \ \mathbf{b}_1 : \quad \overline{\mathbf{y}}_1(\mathbf{b}_1) \in \{\overline{\mathbf{y}}(\mathbf{b}_1), \underline{\mathbf{y}}(\mathbf{b}_1)\}$$

Proof of Lemma 2: If output  $y_1$  belongs to the support of the high productivity firm strategy, from Bayes rule, the CP's posterior beliefs  $v_2$  must be equal to zero unless  $y_1 = y(b_1)$ . Assume  $y_1 \neq y(b_1)$ . Since the CP knows that the firm has a high productivity with certainty, the latter does not obtain a positive second period surplus. Since, whatever the CP's beliefs, the high productivity firm can guarantee itself a non-negative profit in the second period, switching from any  $y_1$  in the support of its strategy and different from both  $\vec{y}(b_1)$  and  $y(b_1)$  to  $\vec{y}(b_1)$  would give such a firm a strictly higher profit. This shows that the support can only consist of two values, as asserted in Lemma 2.

From Lemmas 1 and 2, it follows that there can only be three kinds of equilibria:

- 1) Pooling Equilibrium: Both types of firm choose  $\underline{y}(b_1)$ . The CP's posterior belief is then:  $v_2(a_1,b_1,\underline{y}(b_1)) = v_1$ .
- 2) Separating Equilibrium: The low productivity firm produces  $\underline{y}(b_1)$  and the high productivity firm  $\overline{y}(b_1)$ . Then:  $v_2(a_1,b_1,\underline{y}(b_1)) = 1$  and  $v_2(a_1,b_1,\overline{y}(b_1)) = 0$ .
- 3) Semi-Separating Equilibrium: The low productivity firm produces  $\underline{y}(b_1)$  and the high productivity firm randomizes between  $\underline{y}(b_1)$  and  $\overline{y}(b_1)$ . Then:  $v_2(a_1,b_1,\underline{y}(b_1))\varepsilon(v_1,1)$  and  $v_2(a_1,b_1,\overline{y}(b_1))=0$ .

The determination of the equilibrium type relies on the consideration of the (first period) cost of concealing productivity for a high productivity firm when first period bonus is  $b_1$ . It is denoted  $\Delta(b_1)$ .

$$(4.3) \qquad \Delta(b_1) \equiv b_1[\overline{y}(b_1) - \underline{y}(b_1)] - [\overline{\psi}(\overline{y}(b_1)) - \overline{\psi}(\underline{y}(b_1))] .$$

Note that  $\Delta(0) = 0$  and that  $\Delta(b_1) > 0$  for  $b_1 > 0$ .

Clearly, this short-run cost of concealment has to be compared with the long-run gain. The latter relates with  $\delta \pi^F(\nu)$  which, in the notation of Section 3, will be the discounted surplus of the second period high productivity firm when the CP's a posteriori beliefs are  $\nu$ .

The interest of the above comparison is precisely confirmed by the next proposition.

Proposition 4:  $\forall b_1$ , there exists a unique continuation equilibrium.

If  $b_1$  is such that:  $\Delta(b_1) < \delta \overline{\pi}^F(\nu_1) \quad \text{the equilibrium is pooling}$   $\delta \overline{\pi}^F(1) > \Delta(b_1) > \delta \overline{\pi}^F(\nu_1) \quad \text{the equilibrium is semi-separating}$   $\Delta(b_1) > \delta \overline{\pi}^F(1) \quad \text{the equilibrium is separating.}$ 

Proof: Consider successively the three possible types of
equilibria:

Pooling Equilibrium: In a pooling equilibrium, the long-run loss of playing  $\bar{y}(b_1)$  instead of  $\underline{y}(b_1)$  is  $\delta \bar{\pi}^F(\nu_1)$  since the latter is the high productivity firm's discounted second period surplus when the CP does not acquire information in the first period. Thus a necessary condition for the existence of a pooling equilibrium is:

$$\delta \overline{\pi}^{F}(v_{1}) > \Delta(b_{1})$$

Conversely, let us show that if (4.4) holds, there exists a pooling equilibrium. For that, we must complete equilibrium actions and beliefs with out-of-equilibrium ones. Since the second period strategies of the firm and the CP are uniquely defined in Section 3 for a given  $v_2$ , we just have to describe the CP's beliefs for out-of-equilibrium outputs. There are several ways to choose these beliefs so that the firms actually want to pool at  $\underline{y}(b_1)$ . The simplest way to do so is to take "optimistic" (for the CP) beliefs:  $\underline{v}_1 \neq \underline{y}(b_1)$ :  $v_2 = 0$ , i.e., the CP believes that the firm which chooses  $y_1 \neq \underline{y}(b_1)$  has a high productivity; the corresponding second period scheme has slope s

and lump-sum transfer  $a_2$  such that the high productivity firm makes a zero surplus  $(a_2 = \overline{\psi}(\overline{y}(s)) - s\overline{y}(s))$ . Now it is clear that the efficient firm does not want to reveal its productivity, since the short-run gain to doing so does not exceed  $\Delta(b_1)$  and the long-run loss equals  $\delta \overline{\pi}^F(v_1)$ .

We conclude that there exists a pooling equilibrium if and only if (4.4) is satisfied, and that this equilibrium is unique. The firm, whatever its type, produces  $\underline{y}(b_1)$  in the first period; in the second period the CP chooses a bonus  $b_2 = b(v_1)$  and a lump-sum transfer such that the low productivity firm has a zero surplus. The firm then produces its optimal static output given  $b_2$ .

### Separating Equilibrium

In a separating equilibrium, the high productivity firm produces  $\bar{y}(b_1)$  in the first period and makes a zero surplus in the second. If instead it decided to produce  $\underline{y}(b_1)$  and pool with the low productivity firm, it would lose  $\Delta(b_1)$  in the first period and gain  $\delta \bar{\pi}^F(1)$ , as the CP is convinced that he faces firm  $\underline{\theta}$  when he observes  $\underline{y}(b_1)$ . Thus a necessary condition for the existence of a separating equilibrium is:

$$\delta \overline{\pi}^{F}(1) < \Delta(b_{1})$$

Conversely, it is easy to show that if (4.5) holds, there exists a unique separating equilibrium path that can be supported by out-of-equilibrium beliefs and strategies (take optimistic beliefs for output that differs from  $\underline{y}(b_1)$  and  $\overline{y}(b_1)$ ). Along this equilibrium path, the firm, whatever its type, produces its static optimum in the first

period, the CP then has complete information and chooses  $b_2 = s$  and  $a_2$  so as to extract the whole surplus from the firm. The firm then produces its static optimum for bonus s.

### Semi-Separating Equilibrium:

In a semi-separating equilibrium, the efficient firm is indifferent between revealing its productivity by producing  $\bar{y}(b_1)$ , and pooling with the low cost firm by producing  $\underline{y}(b_1)$ . The short-run loss of pooling is  $\Delta(b_1)$ . The long-run gain is  $\delta \bar{\pi}^F(v_2)$  where  $v_2$  is the CP's second period belief when he observes  $\underline{y}(b_1)$ . Note that  $v_2$  must belong to  $(v_1,1)$  since the low-productivity firm produces  $\underline{y}(b_1)$  with probability one. Thus a necessary condition for existence of a semi-separating equilibrium is that there exists  $v_2$  in  $(v_1,1)$  such that.

$$\delta \bar{\pi}^{F}(\nu_{2}) = \Delta(\nu_{1})$$

From Proposition 3,  $\bar{\pi}^F(v_2)$  is continuous and strictly increasing in  $v_2$ , thus  $b_1$  must satisfy:

$$\delta \overline{\pi}^{F}(\nu_{1}) < \Delta(b_{1}) < \delta \overline{\pi}^{F}(1)$$

Conversely, assume that (4.7) holds. Again from Proposition 3, there exists a unique  $v_2$  in  $(v_1.1)$  such that (4.6) is satisfied. And therefore there exists a unique real number  $x_1$  in (0.1) such that, if the efficient firm produces  $\vec{y}(b_1)$  with probability  $x_1$  and  $y(b_1)$  with probability  $(1-x_1)$ , the CP's posterior belief when he observes  $y(b_1)$ ,  $v_2 = v_1/(v_1 + (1-x_1)(1-v_1))$ , satisfies (4.6). Again it is easy to construct supporting beliefs and strategies

out-of-equilibrium (take optimistic beliefs for outputs that differ from  $\bar{y}(b_1)$  and  $\bar{y}(b_1)$ ).

To conclude, if  $b_1$  satisfies (4.7), there exists a unique semiseparating equilibrium path. The low productivity firm produces  $\underline{y}(b_1)$  and the high productivity firm randomizes between  $\underline{y}(b_1)$  and  $\overline{y}(b_1)$ . When observing the non-fully revealing output  $\underline{y}(b_1)$ , the CP chooses a bonus  $b_2 = b(v_2)$  (where  $v_2$  is defined by (4.6)) and  $a_2$  so as to extract the low productivity firm's surplus. When observing  $\overline{y}(b_1)$ , he chooses bonus s and  $a_2$  so as to extract the high productivity firm's surplus.

To finish the proof of Proposition 4, we observe that for a given  $b_1$ , one of the three mutually exclusive conditions (4.4), (4.5) and (4.7) must hold.

Q.E.D.

The continuation equilibrium depends on the first period cost of concealing productivity  $\Delta(b_1)$ , that is represented in Figure 3.

We shall say that  $\Delta(b_1)$  is "well-behaved" if it increases with  $b_1$  (Figure 3.1,; this means that when the bonus increases, the high productivity firm must incur higher and higher losses in order to mimic the low productivity firm. For example, one can easily show that if the cost function is quadratic in output,  $\Delta(b_1)$  is increasing and convex. For a "well-behaved" function, a higher bonus leads to more separation. However  $\Delta(b_1)$  need not be increasing in  $b_1$ . Its shape depends on third derivatives of the cost function, on which we have little information. This makes the study of the optimal first-period incentive scheme analytically complicated.

 $\frac{\text{Figure 3}}{\text{Regions for continuation equilibrium}}$  (P = pooling, SS = semi-separating, S = separating)

Figure 3.1

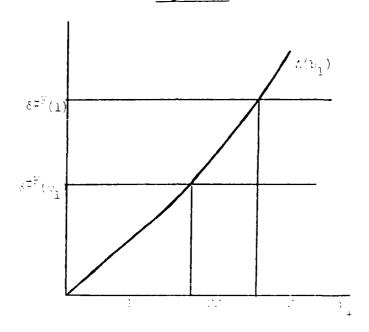
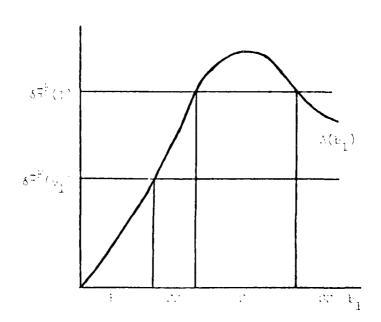


Figure 3.7



### c) Strongly Selective Schemes:

Until now we have assumed that the CP did not want to choose incentive schemes such that the low productivity firm does not operate. This assumption is easily justified for the second period scheme: the low productivity firm produces  $\underline{y}(b_1)$  in the first period and when observing  $\underline{y}(b_1)$ , the CP has belief  $v_2 > v_1$ . Thus if in a one-period situation, the CP wants both firms to operate, a fortiori he does so in the second period when he observes  $\underline{y}(b_1)$ . What about the first-period scheme? It is easy to check that a scheme such that no firm operates cannot be optimal (it is disastrous from the first period point of view, and does not bring any information). Assume now that the first-period scheme is such that only the efficient firm produces. It must be the case that:

$$(4.8)$$
  $a_1 + b_1 \underline{y}(b_1) - \underline{\psi}(\underline{y}(b_1)) < 0$ 

and

(4.9) 
$$a_1 + b_1 \bar{y}(b_1) - \bar{\psi}(\bar{y}(b_1)) > \delta \bar{\pi}^F(1)$$

since the efficient firm can always pretend it has low productivity by not producing. This type of scheme can be called a <u>strongly selective</u> scheme since in order to induce revelation it forces one type of firm not to operate. The idea behind proposing a strongly selective scheme is to relax the self-selection constraint: the high productivity firm is less tempted not to produce than to produce  $\underline{y}(b_1)$ . In general we cannot exclude on a priori grounds these strongly selective schemes, and

in order to solve the CP's problem, we must compare the best such scheme (if these schemes exist) and the best scheme obtained under the assumption that both firms operate

In the previous discussion, we assumed that a firm that does not produce in the first period is able to produce in the second. One can easily imagine circumstances - e.g., managers leave - under which this does not hold. Thus an alternative assumption is that only firms that produce in the first period can produce in the second. Under this assumption, also a strongly selective scheme cannot be ruled out; indeed an example is given in Section 5(c) in which the best first-period scheme leaves the inefficient firm out ("for ever"), in spite of the fact that in a static framework the CP would want both firms to operate.

Not to multiply cases, we shall assume in the next section that strongly selective schemes are not optimal.

# 5. Long-Run Relationship and the Ratchet Effect: The Optimal Dynamic Policy

# a) The CP's first-period problem: key facts:

Although it is needed for a comprehensive study of the phenomenon under consideration, this technical section can be skipped by the reader who is not interested in the detailed proof. The results as well as comments are presented in next subsection.

The problem we face now is to find the CP's optimal first-period decision, i.e., the optimal  $b_1$  (remember that if we rule out strongly selective schemes,  $a_1$  is a function of  $b_1$ ). The existence of a unique continuation equilibrium for any given  $b_1$  makes the CP's first-

period maximization problem well-defined. We were not able to prove that in general the solution to this problem is unique; however, a multiplicity of first-period optimal schemes is not too severe a problem, since, once the CP has chosen  $(a_1,b_1)$ , the continuation equilibrium is unique (there is no problem of coordination between equilibria). The welfare outcome W is also unique.

One important issue, on which one would expect clarifications from the analysis, concerns the relative magnitude of the optimal first-period bonus in a dynamic context and the static optimal bonus. One is tempted to believe a priori that the Planner should be more "generous" in a dynamic context in order to obtain information. In fact, the Planner problem is rather complex and in order to understand it, we have to single out a certain number of key facts which are gathered in the next lemmas.

A preliminary lemma asserts that more information is indeed desirable to the Center.

Lemma 3: The expectation of the optimal second period aggregate social welfare  $W_2$ , for a given first-period scheme, increases with the number of revealing high productivity firms  $x_1$  (or decreases with the a posteriori probability that a pooling firm is of high productivity).

The argument goes as follows: For the firms that switch from pooling in the reference situation to revealing, second-period social welfare is increased since the second-period scheme is now the full information scheme  $(a_2 = \overline{\psi}(\overline{y}(s)) - s\overline{y}(s), b_2 = s)$ . For the pooling firms, the CP can always duplicate his previous second-period incomplete

information scheme in the reference situation (which is now sub-optimal) and leave social welfare unaffected.

We now consider the effect of the first-period choice between pooling, semi-separating and separating on first-period aggregate welfare.

First, let us assume that firms are pooling. First-period aggregate welfare is:

$$w^{P}(b_{1}) = p\underline{y}(b_{1}) - \lambda(\underline{a}(b_{1}) + b_{1}\underline{y}(b_{1})) - v_{1}\underline{\psi}(\underline{y}(b_{1})) - (1 - v_{1})\overline{\psi}(\underline{y}(b_{1}))$$

where  $\underline{a}(b_1)$  is given by  $\underline{a}(b_1) + b_1 \underline{y}(b_1) - \underline{\psi}(\underline{y}(b_1)) = 0$ We prove:

Lemma  $\mu$ : For  $b_1 \leq s$ ,  $W^P(b_1)$  increases with  $b_1$ .

Proof:

$$\frac{\mathrm{d} \underline{w}^{\mathrm{P}}}{\mathrm{d} b_{1}} = \left[ p - \lambda b_{1} - v_{1} \underline{\psi}' - (1 - v_{1}) \overline{\psi}' \approx \frac{\mathrm{d} \underline{y}}{\mathrm{d} b_{1}} - \lambda \left[ \underline{y}(b_{1}) + \frac{\mathrm{d} \underline{a}}{\mathrm{d} b_{1}} \right] \right]$$

As  $\overline{\psi}'(\underline{y}(b_1)) \le \underline{\psi}'(\underline{y}(b_1)) = b_1$ , it follows from the computation of  $d\underline{a}/db_1$  that:

$$\frac{dW^{P}}{db_{1}} > \left[\bar{p} - (1 + \lambda)b_{1}\right] \frac{dy}{db_{1}}$$
 Q.E.D.

Define:

$$\tilde{W}_{1}(b_{1},x_{1}) = [v_{1} + (1 - v_{1})(1 - x_{1})]W^{P}(b_{1}) + [x_{1}(1 - v_{1})]$$

$$[p\overline{y}(b_{1}) - \lambda(\underline{a}(b_{1}) + b_{1}\overline{y}(b_{1})) - \overline{\psi}'(\overline{y}(b_{1}))]$$

 $\tilde{w}_1(b_1,x_1)$  is the aggregate first period welfare when the first-period bonus is  $b_1$  and the proportion of revealing high productivity firms is  $x_1$ .

Lemma 5: The first-period aggregate welfare  $\tilde{W}_1$  increases with the proportion  $x_1$  of revealing high productivity firms, for any  $b_1 \le s$ .

(Note that Lemmas 4 and 5 hold independently of the value of the parameter  $\nu_{_{1}})\,.$ 

Proof: Assume that for  $b_1 \le s$ , one high productivity firm switches from pooling to separating. This move increases the first-period profit of the firm. But it also increases the CP's "social concern" utility for the first period: social concern indifference curves are straight lines of slope s (see Figure 2) and separation moves the firm's bundle to the right along a line of slope  $b_1$ , i.e., on a higher social concern indifference curve when  $b_1 \le s$ . Q.E.D.

The following corollary is a straightforward implication of Lemmas 3 and 5.

Corollary 5.1: For any  $b_1 \le s$ , intertemporal aggregate social welfare increases with the number of revealing high productivity firms.

We now show that if  $b_1 > s$ ,  $\tilde{W}_1(s,1) > \tilde{W}_1(b_1,x_1)$ ,  $x_1 \neq 1$ .

Lemma 6: A separating equilibrium with bonus s would be superior from the first-period aggregate welfare point of view to a

semi-separating (or pooling) equilibrium with  $b_1 > s$ .

<u>Proof</u>: Consider the reference situation, a semi-separating or pooling equilibrium with  $b_1 > s$  and compare it to a situation of separation occurring with  $b_1' = s$  and  $a_1' = \underline{a}(b_1) + b_1\underline{y}(b_1) - s\underline{y}(b_1)$  (the new budget line passes through  $(\underline{y}(b_1), \underline{a}(b_1) + b_1\underline{y}(b_1))$ ).

In this new situation, consider successively:

- The first-period aggregate welfare restricted to low productivity firms: "social concern" first-period utility remains constant but the individual profit increases when the firm reacts to the new offer.
- The same argument holds for the first-period aggregate welfare restricted to high productivity firms which were pooling in the initial situations.
- For high productivity firms which separate in the initial situation, the total effect can be decomposed (as in the proof of Proposition 1) into two parts: a substitution effect along the initial indifference curve which by definition does not affect the firm's utility and increases the (first-period) social concern utility; an income effect which is negative from the firm's point of view but positive in terms of aggregate welfare.

Finally, the impact on first-period aggregate welfare of switching to the new scheme is unambiguously positive.

Still assuming separation, we can now switch to a scheme with  $b_1 = s$ , bringing the low productivity firm to its minimum utility, a change which still increases first-period aggregate welfare (such a change has only positive income effects in terms of aggregate welfare).

Q.E.D.

The preceding facts hold independently of any assumption on  $\Delta(b_1).$  The next property relies on the fact that  $\Delta(b_1)$  increases with  $b_1$ .

Lemma 7: Assume that  $\Delta(b_1)$  is an increasing function of  $b_1$ . Then if  $b_1 < b(v_1)$  induces a semi-separating or pooling continuation equilibrium, it cannot be a first-period optimal dynamic bonus.

<u>Proof:</u> Let  $\bar{x}_1$  denote the proportion of revealing high productivity firms with the scheme  $\bar{b}_1 < b(\nu_1)$ . Consider  $\bar{\nu} = \nu_1/(\nu_1 + \bar{x}_1(1 - \nu_1))$  and  $b(\bar{\nu})$  the static optimal bonus associated with  $\bar{\nu}$ . From Proposition 3,  $\bar{\nu} > \nu_1$  implies  $b(\bar{\nu}) > b(\nu_1)$  (but  $b(\bar{\nu}) < s$ ). (Note that for  $b(\bar{\nu})$  the proportion of high productivity revealing firms  $\bar{x}_1$  is strictly greater than  $\bar{x}_1$ : this follows from the property of  $\Delta(b_1)$ , Proposition 3, and the consideration of formula (4.6).)

Now instead of  $\mathbb{B}_1$ , take  $\mathfrak{b}(\overline{\nu})$  as first-period bonus. The induced change can be decomposed by a thought experience into two parts: - A change from  $\mathbb{B}_1$  to  $\mathfrak{b}(\overline{\nu})$  leaving the number of high productivity revealing firms unaffected. For the group of low productivity firms and revealing high productivity firms, this change, by definition of  $\mathfrak{b}(\overline{\nu})$ , is favorable from the CP's viewpoint in the first period and nothing is changed in the second period. For the subgroup of non-revealing high productivity firms, the move is also socially farorable in the first period (use Lemma 4 for  $\nu_1=0$ ) and indifferent in the second. So, the intertemporal aggregate welfare increases. - A change at  $b_1 = b(\overline{\nu})$  in the proportion of revealing high productivity firms. This change is favorable to the CP in the first period (Lemma 5) and in the second period (Lemma 3).

The conclusion follows.

Q.E.D.

### b) The Optimal First-Period Bonus: Results:

We have now enough understanding of the factors governing the CP's intertemporal welfare to be able to single out conclusions on the optimal first-period dynamic bonus.

First, the fact that the static optimal bonus is smaller than the social value of the good, has a somewhat weaker dynamic counterpart:

Proposition 5: If the continuation equilibrium associated with the bonus s is separating, then  $b_1^D$ , the optimal dynamic first-period bonus, satisfies

$$b_1^D \leq s$$
.

 $\underline{Proof}$ : Suppose that  $b_1^D > s$ . Then:

- either  $b_1^D$  is separating, a fact ruled out by Proposition 2(i),
- or  $b_1^D$  is not separating but from Lemma 6, the scheme with bonus s would be better from the CP's first-period point of view. As it is also better for the second-period aggregate welfare (this is a variant of Lemma 3), a contradiction obtains. Q.E.D.

Let us now derive a lower bound for  $\mathfrak{b}_1^{\mathbb{D}}$ . The following is not very surprising:

Proposition 6: If the continuation equilibrium associated with  $b_1 = b(v_1)$  is separating, then:

$$b_1^D = b(v_1)$$

<u>Proof</u>: Assume the contrary:  $\overline{b}_1 \neq b_1(v_1)$  is optimal.

- $\bar{b}_1$  cannot be separating: given that the informational contents of  $\bar{b}_1$  and  $b(\nu_1)$  are the same, this would contradict the definition of  $b(\nu_1)$ .
- $\delta_1$  cannot be a pooling or semi-separating equilibrium with  $\delta_1 < s$ . According to Lemma 5, a separating equilibrium with  $b_1 = \delta_1$  would be superior for the first-period aggregate welfare to the equilibrium under consideration, it is also superior for the second-period point of view. However, this hypothetical separating equilibrium associated with  $b_1 = \delta_1$  is itself dominated by the scheme  $b_1 = b(\nu_1)$ , a contradiction.
- $b_1$  cannot be a pooling or semi-separating equilibrium with  $b_1 > s$ .

  According to Lemma 6, such an equilibrium would be dominated by a hypothetical separating equilibrium with  $b_1 = s$ , which in turn is dominated by  $b_1 = b(v_1)$ , a contradiction. Q.E.D.

Proposition 6 looks intuitively reasonable. If the second best static bonus induces in addition full revelation in a dynamic context, it is the dynamic optimum. What is surprising is rather that the proof is more intricate than expected. The reason is that pooling in the first period (or semi-separation), a possibility which is not open in the second best static problem, may be an interesting option from the

viewpoint of the first-period "social concern". The fact that it is not so favorable after all - the argument on which the proof relies - is a consequence of Lemmas 5 and 6.

In Proposition 7, below, we will see that another intuitively appealing assertion, that the dynamic optimal first-period bonus is higher than the second best optimal static bonus, holds only in favorable cases and for the same reason as above requires a non-straightforward proof. But beforehand we will give a corollary to Proposition 6:

Corollary:  $\exists \delta_0$  such that for  $\delta \leq \delta_0$ , the first-period dynamic optimum is  $b(\nu_1)$ .

Proof: Take  $\delta_0$  small enough such that  $\Delta(b(\nu_1)) > \delta_0 \overline{\pi}^F(1)$ . Hence  $b(\nu_1)$  is necessarily in the separating zone and Proposition 6 applies. Note however that  $\delta_0 < 1$ ; for  $\delta = 1$ ,  $b(\nu_1)$  cannot be in the separating zone (remember that  $\overline{\pi}^F(1) = \Delta(s)$ ).

Restricting now our attention to the well-behaved case, i.e., the case in which the incentive to conceal one's information decreases with the bonus, we have:

Proposition 7: In the well-behaved case, the optimal first-period bonus exceeds the optimal one-period bonus:  $b_1^D > b(v_1)$ .

<u>Proof</u>: Suppose  $b_1^D \le b(v_1)$ . According to Lemma 4 and because of  $b(v_1) < s$ ,  $b_1$  cannot be pooling without being at the frontier of semi-separation. But because of Lemma 7,  $b_1$  cannot be semi-separating

(weakly or strictly). Hence, if  $\ b_1^D$  is not semi-separating it is greater than  $\ b(\nu_1^{}).$ 

From Proposition 6, if  $b_1^D$  is separating, it equals  $b(\nu_1)$ . The conclusion follows. Q.E.D.

c) An Example: Let us assume that the cost function has the following simple form:  $\psi(y,\theta)=ky^2/2\theta$  (where k>0). The reader can easily check that for this cost function:

$$y(b,\theta) = \frac{b\theta}{k}$$

(5.2) 
$$\Delta(b) = \frac{b^2}{2k\overline{\theta}} [\overline{\theta} - \underline{\theta}]^2$$

(5.3) 
$$b(v) = s \frac{v\theta + (1 - v)\theta}{v\theta + (1 - v)\theta[1 + \frac{\lambda}{1 + \lambda} \frac{\overline{\theta} - \underline{\theta}}{\overline{\theta}}]}$$

(5.4) 
$$\overline{\pi}^{F}(v) = \frac{b^{2}(v)}{2k} (\overline{\theta} - \underline{\theta})$$

 $\Delta(b)$  is represented in Figure 4. Let  $\hat{b}$  and  $\hat{b}$  be the bonuses that mark the upper limits of the pooling and semi-separating regions. From (5.2) and (5.4) and Section (b),  $\hat{b}$  and  $\hat{b}$  are defined by:

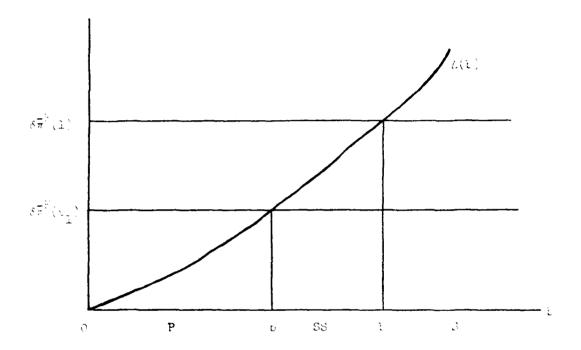
$$(5.5) \qquad \qquad \hat{b} = \mu b(\nu_1)$$

(5.6) 
$$\hat{b} = \mu s$$

where

$$\mu \equiv \sqrt{\frac{\delta}{\frac{\overline{\theta} - \underline{\theta}}{\overline{A}}}}$$

Figure 4



Note that if  $\mu < b(\nu_1)/s$ , then  $b(\nu_1)$  belongs to the separating region and thus is optimal. This case arises when the future is not too highly valued and the productivity differential is important. On the contrary, if  $\mu$  is high,  $\hat{b}$  and  $\hat{b}$  are very high relative to s, and it becomes very costly to separate the two types of firms. Moreover, the second-period social gain of having better information is very small

if the firms are very alike. Thus the optimal bonus b<sub>1</sub> leads to pooling, and from Lemma 3, is greater than s. It is also possible to show that, for some values of the parameters, it is socially optimal to induce semi-separation.

Consider now the possibility of using strongly selective schemes. We shall assume that a firm that does not produce in the first period is able to produce in the second. It is clear that for a strongly selective scheme to be optimal, the lump-sum al must make (4.9) binding:

(5.8) 
$$a_1 = \delta \bar{\pi}^F(1) + \bar{\psi}(\bar{y}(b_1)) - b_1 \bar{y}(b_1)$$

Let us look for the set of bonuses such that the scheme is actually strongly selective, i.e., (4.8) is satisfied. From a simple computation, this is the case if and only if:

$$(5.9) b1 > \sqrt{\delta}s$$

We shall assume that  $\delta$  is less than 1. Then clearly the best strongly selective scheme has slope s and gives first-period welfare  $\mathbf{w}_1^{SSS}$ :

$$W_{1}^{SSS} = (1 - v_{1})[p\bar{y}(s) - \bar{\psi}(\bar{y}(s))(1 + \lambda) - \lambda(\delta\bar{\pi}^{F}(1))]$$

$$= \frac{(1 - v_{1})ps\bar{\theta}}{2k} \left[1 - \frac{\lambda\delta}{1 + \lambda} \left(\frac{\bar{\theta} - \theta}{\bar{\theta}}\right)\right]$$

What is the best scheme under the assumption that both firms operate? If  $\mu = \sqrt{\left[\left(\delta/((\overline{\theta}_- - \theta)/\overline{\theta})\right]\right]}$  is "high enough", say above  $\overline{\mu}$ ,

inducing separation or semi-separation implies a very high bonus and a first-period negative welfare, as the reader will check. Thus only a bonus that induces firms to pool can be preferred to strong selection. The best pooling bonus can be found by maximizing  $W^P(b)$ :

$$(5.11) \mathbf{w}^{\mathbf{P}}(b) = \mathbf{p} \underline{\mathbf{y}}(b) - \underline{\mathbf{\psi}}(\underline{\mathbf{y}}(b))(1 - \lambda) + (1 - \nu_1)(\underline{\mathbf{\psi}}(\underline{\mathbf{y}}(b)) - \overline{\mathbf{\psi}}(\underline{\mathbf{y}}(b)))$$

(The second term in (5.11) represents the savings in effort when the firm is efficient.) The optimal pooling bonus is then:

(5.12) 
$$b^{P} = \frac{s}{1 - v_{1} + \theta - \theta}$$

and leads to a first-period well'are level:

(5.13) 
$$W^{p} = \frac{ps \, 9}{2k} - \frac{1}{1 + \lambda} \frac{\theta - \theta}{\pi}$$

Thus  $W_1^{SSS}$  exceeds  $W^P$  if and only if.

$$(5.14) (1 - v_1)(1 - \frac{\delta \lambda}{1 + \lambda} \frac{\theta - \frac{\theta}{1}}{\theta}) (1 - \frac{1 - v_1}{1 + \lambda} \frac{\theta - \frac{\theta}{1}}{\theta}) > (1 - \frac{\theta - \frac{\theta}{1}}{\theta})$$

Now, consider a set of parameters such that:

- 1) (5.14) holds.
- 2) In a static framework, the CP is indifferent between having both firms operate and having only the efficient firm operate.

3)  $\mu$  is "high enough", i.e., separation and semi-separation are not desirable. Tedious considerations show that it is indeed possible to find such parameters. 10/

We claim that the best scheme is then the best strongly selective scheme. If the CP chooses a strongly selective scheme, the second period welfare is nothing but the full information one-period welfare. If the CP chooses to have both firms pool, the second period welfare is the one-period welfare with prior  $\nu_1$ . Since  $W_1^{SSS} > W_1^P$ , the strongly selective scheme dominates the pooling scheme. One can show that Proposition 8 is valid under the two alternative assumptions, i.e., that a firm does not produce in the first period can or cannot produce in the second:

Proposition 8: Even if the CP wants both types of firms to operate in a static context, his best strategy in a dynamic framework may result in strong selection, i.e., elimination of the less efficient firm.

Proposition 8 shows that the elimination of "lame ducks", which in a dynamic model induces more revelation from the efficient firm, may be socially desirable even if it is not in a traditional static planning model.

d) Non-Linear Schemes: In this section we want to show that the main ideas developed for linear schemes carry over to optimal non-linear schemes. Again we will confine ourselves to the two-firm case. More general results for arbitrary distributions are still to be found.

The CP chooses two pairs of incentives schemes  $S_1 = \{(\underline{y}_1, \underline{R}_1), (\overline{y}_1, \overline{R})\}$  in the first period and  $S_2 = \{(\underline{y}_2, \underline{R}_2), (\overline{y}_2, \overline{R}_2)\}$  in the second. The timing is the same as before. We will use the following proposition and lemmas:

Proposition 4': For a given  $S_1$ , there exists a unique continuation equilibrium.

<u>Proof:</u> We just sketch the proof since it is very similar to that of Proposition 4. Let  $\Delta$  be the cost for the high productivity firm to concealing its productivity:

$$\Delta \equiv \left[ \overline{R}_1 - \overline{\Psi}(\overline{y}_1) \right] - \left[ \left( \underline{R}_1 - \overline{\Psi}(\underline{y}_1) \right) \right]$$

Let  $v_2$  be the CP's second period beliefs,  $\overline{\pi}_{NL}^F(v)$  be the efficient firm's one-period surplus when the CP has beliefs v, and  $\underline{y}^v$  be the output chosen by the CP with beliefs v for firm  $\underline{\theta}$ . It is easy to show that  $\overline{\pi}_{NL}^F(v)$  increases with v and that  $\overline{\pi}_{NL}^F(1) = \underline{\psi}(\underline{y}(s)) - \overline{\psi}(\underline{y}(s))$ .

Thus if  $\Delta > \delta \overline{\pi}_{NL}^F(1)$  the efficient firm reveals its productivity; if  $\Delta < \delta \overline{\pi}_{NL}^F(v_1)$ , it chooses  $(\underline{y}_1,\underline{R}_1)$ ; and if  $\delta \overline{\pi}_{NL}^F(v_1) < \Delta < \delta \overline{\pi}_{NL}^F(1)$ , the continuation equilibrium is semi-separating. Q.E.D.

Let us now consider the CP's first period decision problem. We shall assume that the CP chooses to have both firms operate. Much of the discussion of Sections h(b) and 5(c) on strongly selective schemes would carry over so non-linear schemes. We first prove the following two lemmas.

<u>Lemma 8</u>:  $R_1 = \psi(y_1)$ 

Proof: It suffices to notice that the continuation equilibrium depends only on  $\Delta$  and not directly on  $R_{-1}$  (as long as both firms operate). Q.E.D.

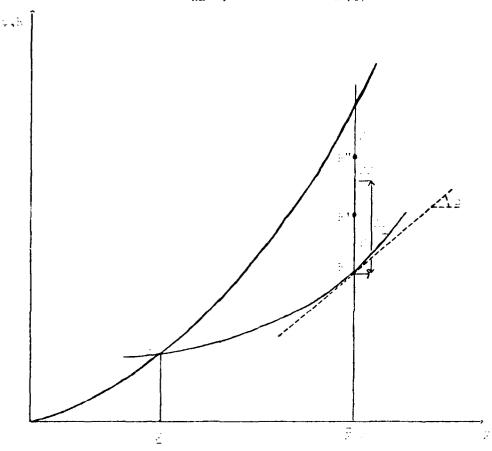
 $\underline{\text{Lemma 9}} : \quad \bar{y}_1 = \bar{y}(s)$ 

<u>Proof:</u> Let  $P' = (\overline{y}_1, R_1)$  and consider the indifference curve  $\overline{\pi}^F$ , through P'. If  $\overline{\Psi}'(\overline{y}_1) \neq s$ , then the CP can do better by proposing instead of P' the point P where the tangent to the indifference curve through P' is of slope s, (unless  $\Delta$  induces firms to pool, in which case P' is irrelevant). From P' to P,  $\Delta$  and  $\overline{\pi}^F$  remain the same and  $\pi^{CP}$  increases.

Using Lemmas  $\theta$  and  $\theta$ , we can represent the CP's first period decision problem with the following diagram:

Figure 5

$$RR' = \delta \overline{\pi}_{NL}^{F}(v_{1})$$
,  $RR'' = \delta \overline{\pi}_{NL(1)}^{F}$ 



We see that the CP's choice amounts to that of an output for firm  $\theta$ :  $\underline{y}_1$  and of a cost of concealment for firm  $\theta$ . Note that the optimal choice of  $\Delta$  given  $\underline{y}_1$  will either result in a pooling equilibrium or belong to the interval  $\left[\delta\overline{\pi}_{NL}^F(\nu_1),\,\delta\overline{\pi}_{NL}^F(1)\right]$  (resulting in a separating or semi-separating equilibrium): Among the  $\Delta$  that induce a revealing equilibrium, the CP always prefers the lowest:  $\delta\overline{\pi}_{NL}^F(1)$ . We can now prove Proposition 6':

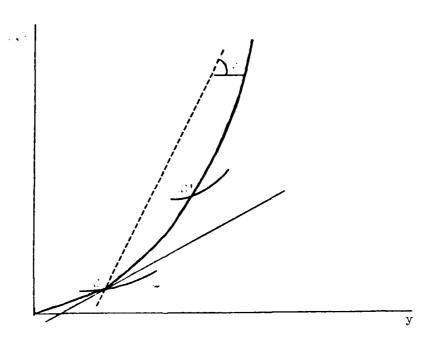
<u>Proposition 6'</u>: For the optimal first period non-linear scheme, the marginal cost of the low productivity firm  $(\underline{\psi}'(\underline{y}_1))$  exceeds that obtained for the one-period optimal non-linear scheme  $(\underline{\psi}'(\underline{y}(\nu_1)))$ .

### Proof:

- (S) First assume that the CP induces a revealing equilibrium  $(\Delta = \delta \overline{\pi}_{NL}^F(1)). \quad \text{Then} \quad \underline{y}_1 = \underline{y}(\nu_1) \text{ since the CP's optimization problem is identical with the one-period problem with beliefs } \nu_1, \text{ up to the additional separation cost } (1 \nu_1) \{\lambda \Delta\}.$
- (P) Next assume that the CP decides to have both firms pool (by choosing e.g.,  $\Delta$  = 0). We claim that  $\underline{\psi}'(\underline{y}_1)$  > s. Figure 6 is drawn under the assumption that  $\underline{\psi}'(\underline{y}_1)$  < s:

Figure 6

Pooling equilibrium  $\underline{\psi}'(\underline{y}_1) < s$ 



In Figure 6, both firms pool at S. Clearly a point such as S' would be preferred by the CP and the high productivity firm and leaves the low productivity firm indifferent  $\frac{11}{}$ 

(SS) Lastly assume that the equilibrium is semi-separating. Let us show that if  $\psi'(\underline{y}_1) < \psi'(\underline{y}(\nu_1))$  the CP can induce a semi-separating equilibrium with a higher welfare level. Indeed assume that he chooses  $\underline{y}(\nu_1)$  instead of  $\underline{y}_1$  and that he maintains the same  $\Delta$ . From Proposition 4', the proportion  $x_1$  of high productivity firms that reveal their information remains the same. Now perform a statistical decomposition into two populations: in the first population, put all revealing high productivity firms and a fraction of low productivity firms such that the proportion low/high productivity remains the same as the initial one; put the other firms in the second population. For the first population, the optimum output for the low productivity firm is  $\underline{y}(\nu_1)$ . For the second population, welfare is higher for  $\underline{y}(\nu_1)$  than for  $\underline{y}_1$  from (P). Thus  $\underline{y}_1 > \underline{y}(\nu_1)$ . Q.E.D.

### 6. Alternative Incentive Schemes

In this paper we considered only optimal linear and non-linear schemes. There has recently been a lot of discussion about intermediate schemes, in particular the so-called old and new Soviet incentive schemes (see, e.g., Fan [1975], Bonin [1976], Ekern [1979] and Holmstrom [1979]).

We shall represent the old and new Soviet incentive schemes as piecewise linear schemes with respectively one and two kinks, such that

the bonus (slope) decreases at the kink(s). These two schemes are pictured in Figures 7 and 8.

Figure 7
Old Soviet Incentive Scheme

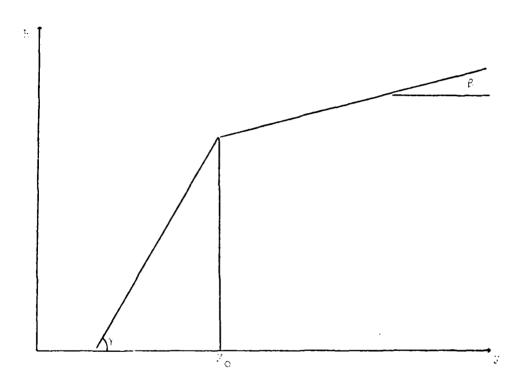
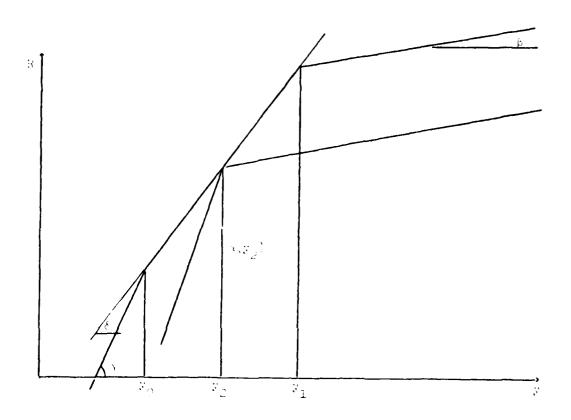


Figure 8

New Soviet Incentive Scheme



In the old scheme, the CP chooses  $\{y_0,\alpha,\beta,\gamma\}$  with  $y_0>0$ ,  $\gamma>\beta>0$  and rewards the firm:  $[\alpha+\beta(y-y_0)]$  if  $y>y_0$  and  $[\alpha-\gamma(y_0-y)]$  if  $y< y_0$ . In the new scheme, the CP chooses  $\{y_0,y_1,\alpha_1,\beta,\gamma,\delta\}$  with  $y_0>0$ ,  $y_1>0$ ,  $\gamma>\delta>\beta>0$ ,  $\alpha_0+\delta(y_1-y_0)=\alpha_1$  and rewards the firm  $[\alpha_0-\gamma(y_0-y)]$  if  $y< y_0$ ,  $[\alpha_0+\delta(y-y_0)]$  if  $y_1>y_0$ , and  $[\alpha_1+\beta(y-y_1)]$  if  $y>y_1$ . This, of course, is not the usual interpretation of the new Soviet incentive scheme. In the usual formalization, the CP chooses the line with slope  $\delta$ , the slopes  $\beta$  and

 $\gamma$ , and two bounds  $y_0$  and  $y_1$   $(y_0 < y_1)$ ; then the firm chooses a target  $y_2$  within  $(y_0,y_1)$  and is rewarded according to the associated old scheme  $\{y_2,\alpha(y_2),\beta,\gamma\}$  where  $\alpha(y_2)$  is the revenue corresponding to  $y_2$  on the line with slope  $\delta$  (see Figure 8). Similarly, it is possible to interpret the old scheme as the announcement by the firm of a target which is restricted by only one constraint. The announcement interpretation is not relevant to our situation, in which the social price of output is given. However, if p is not given, the announcement effect may matter. For example, it may allow the CP to develop or look for a substitute supply of the good earlier than would otherwise be permitted.  $\frac{12}{\epsilon}$ .

We just considered one kind of uncertainty faced by the CP, namely that arising from incomplete information. Assume now that the firm's output is random; and that the social price of output p is not given. One can, for example, imagine that the firm produces an intermediate good for which no close substitute exists. It is in general important that the firm's production does not fall under some threshold in order not to create a severe shortage of the good. On the other hand, a very high production is a relatively minor improvement over normal production since downstream firms then form a bottleneck. It is thus worth encouraging managers and workers to work very hard in (industry-specific) bad states of nature, and not rewarding them too generously for overproduction. Non-linear schemes allow the CP to induce the firm to internalize to some extent the non-linearity of the social value of output.

As we said, the latter point as well as the announcement effect are absent in our model since we assumed that the social price of output is given (for example, determined by the world price). Hence, in our context, the old and new Soviet incentive schemes - which are intermediate between linear and fully non-linear schemes - are of limited interest. Therefore, we will just mention without proofs the following results for the two-firm case.

- 1) Static model:
- (a) The best old scheme is also the best new scheme (this result of course does not generalize to more than two firms).
- (b) The best old scheme dominates the best linear scheme. For the best old scheme, the shadow price of output s exceeds the marginal cost of the low productivity firm, which in turn exceeds the marginal cost of the high productivity firm (this last comparison differs from the usual conclusion for optimal fully non-linear schemes).
- 2) Dynamic model: There is a ratchet effect, i.e., firms may want to hide their information by not choosing their static optimizing output (this result of course holds for most classes of incentive schemes).

An interesting line of research would be the analysis of the relative sensitivity of different classes of schemes to the ratchet effect (for example: is the payoff to pretending to have a low productivity smaller in the new Soviet incentive scheme than for linear schemes?).

### 7. Commitment

We have analyzed the ratchet effect in the context of a game. An other approach to this problem assumes that the CP is able to commit himself to an intertemporal sequence of incentive schemes; in this case he announces the current scheme as well as the revision procedure, and the firm solves its dynamic programming problem given the CP's plans. This is the set-up considered by Weitzman [1980] and Holmstrom [1979]. Holmstrom show that, in spite of the ratchet effect, a revision procedure is preferable to a fixed scheme.

Note that the CP always prefers committing himself. Indeed he can always reach our (unique) equilibrium social welfare by duplicating his optimal strategy(ies). The firm's dynamic programming response is then its optimal strategy in the non-commitment game.

On the other hand, casual empiricism suggests that central planners generally do not commit themselves to revising incentive schemes in a given way. We must then ask why they do so. Although this question is out of the scope of this paper, we would like to mention some elements of answer. One possibility, of course, is that the CP is not aware of the benefits of commitment. But there may be deeper reasons. First, commitment may not be credible. The CP is free to design incentive schemes, but also to change them over time at his discretion. Also the planner may be replaced and the new planner may not feel obliged to abide by the schemes designed by his predecessor (remember that the time period in planning is not short). Second, the costs of designing intertemporal plans may be very high, in particular

when there is a high number of economic units. Third, it is possible that the CP can over time obtain new information about the firm (other than its output) or about the future social price of output (for example, the CP obtains inside information about effort, observes indicators other than output, etc.). This fact by itself does not contradict the general proposition that in the absence of "transaction costs" the CP is better off committing himself if he can. But it greatly increases commitment costs: The CP must then give a detailed account of the future incentive schemes as a function of his information. 13/

### 8. Alternative Interpretations of the Model

The idea that if the relationship between two economic agents is not one-shot, these have incentives to hide their information, is a very general one. For example, the ratchet effect is ubiquitous in all relationships within a firm or institution (a zealous and efficient typist quickly becomes overloaded with work; a complacent employee or manager gets to do the ingrate work, etc.). More generally, it pervades any principal-agent relationship. An example similar to the one presented here is that of an efficient regulated firm which sees the price for its output revised downwards over time. Another example of regulation concerns pollution controls. Yao [1982] considers the dynamics of regulation of automobile emissions control and its effect on the strategy of car manufacturers, also from a non-commitment point of view. The analysis of the dynamics of the principal-agent relationship

could also be relevant to optimal taxation: the government over time learns the ability of tax-payers and in principle might want to use this information to design individual tax-schemes. Why it in general does not do so is an interesting question, on which this kind of analysis may throw some light. Lastly we note that a kind of ratchet effect also arises in bargaining situations (actually this analogy was the starting point of our analysis). Consider two firms or governments bargaining repeatedly over different contracts. Lastly Each partner would like to establish a reputation for being a tough bargainer in order not to face tougher and tougher behavior from the other bargainer.

### Appendix - Proof of Proposition 3

Proposition 3: b(v) and  $\bar{\pi}^F(v)$  are increasing in v

### Proof:

Let us prove that the optimal bonus grows with the proportion of low productivity firms.

Let  $\nu'$  be greater than  $\nu$  and suppose that  $b(\nu') < b(\nu)$ . Decompose the "population"  $\{\nu', 1 = \nu'\}$  of low and high productivity firms into two subpopulations.

The first one has all high productivity firms and v'' = (v/(1-v))(1-v') low productivity firms.

The second has (v' - v'') = low productivity firms (with <math>v' - v'' = (1/(1 - v))(v' - v) > 0).

For the first population, which has the same relative composition as the economy  $(\nu,\ 1-\nu),\ b(\nu)$  is better from the welfare viewpoint than  $b(\nu').$ 

For the second population of low productivity firms, welfare increases with b, when b < s; hence b(v) is also retter than b(v'). These remarks contradict the fact that b(v') + b(v').

To show that  $\overline{\pi}^F$  is increasing in  $\nu$ , it is eneugh to remark that from the firm's first order conditions  $d\overline{\pi}^F/db = \overline{y}(b) = y(b) < \gamma$ .

Ų.E.D.

### Footnotes

- The formalization of the firm as a unique aggregate agent is standard in the literature. It seems to be an unavoidable first step in the analysis. Note however that facts suggest that the efficiency of incentive schemes do crucially depend upon the way incentives are channelled within the firm through "delegation" systems.
- An alternative interpretation is that there are a large number of firms and that  $f(\theta)$  describes the proportion of firms of parameter  $\theta$ . Although on many points the following analysis remains valid with such an interpretation, the comments made rely on the one-firm assumption.
- 3/ Non-linear schemes with a more traditional interpretation for a quota are discussed in Nection 6.
- Let us briefly consider the case when the distribution of firms is continuous, the continuous density over  $\begin{bmatrix} \theta_{\min}, \theta_{\max} \end{bmatrix}$  being  $f(\theta)$ . We can demonstrate that Proposition 2 is robust to the distribution of firms. There is a slight complication added: One can in general not assume that the CP wants the lowest productivity firm  $\begin{pmatrix} e_{\min} \end{pmatrix}$  to operate. Let  $\theta_0(a,b) > \theta_{\min} \end{pmatrix}$  be the "cut-off productivity", i.e., the productivity under which the firm does not want to operate:  $\max_{y} \{a + by \psi(y,\theta_0(a,b))\} = 0$ . The CP's behavior is described by the following program:

(3.1') 
$$\max_{\{a,b\}} \frac{\int_{0}^{a} \max_{\{a,b\}} \{py(b,\theta) - \lambda(a + by(b,\theta)) - \psi(y(b,\theta),\theta)\} f(\theta) d\theta}{\int_{0}^{a} \max_{\{a,b\}} \{py(b,\theta) - \lambda(a + by(b,\theta)) - \psi(y(b,\theta),\theta)\} f(\theta) d\theta}$$

It is easily shown that the equivalent of (3.3b) is:

(3.3b') 
$$s - b = \frac{1}{1 + \lambda} \frac{\int_{\theta_{O}(a,b)}^{\theta_{max}} \{y(b,\theta) - y(b,\theta_{O}(a,b))\} f(\theta) d\theta}{\int_{\theta_{O}(a,b)}^{\theta_{max}} \frac{\partial y}{\partial b} (b,\theta) f(\theta) d\theta}$$

which demonstrates that s>b. Rather than deriving  $(a,b)^{-1}$ , we observe that the above proof carries over to a continuous distribution. Starting from a bonus scheme (a,b) assumed to be optimal and where the "cut-off firm" is  $\theta$  (a,b), we change it into a bonus scheme (a',s) with the same "cut-off firm". The

above argument corresponding to the case where  $\theta = \theta$  applies here for every  $\theta > \theta$  (a,b), when the firm is of type  $\theta$ , aggregate welfare increases with the new scheme. By summing over  $\theta > \theta$  (a,b), we conclude that  $\theta \in S$ .

- 5/ In fact, it is easy to see that the set of axioms for the principal-agent relationship considered in this latter article, applies here (with slight modifications). Hence the results in this subsection can be viewed as particular cases of the more general propositions of Diesnerie-Beade.
- The fact that  $\delta$  is a common discount factor is not crucia, for the theory although it simplifies the reasoning. However, with different discount factors, the meaning of W, the sum of discounted profit and "separally concerned" we have (see (4.7)), is ambiguous.
- For notational simplicity, we shall use the same letters for strategies and actions.
- Note however, that the insert strately which is adopted does not strictly dominate, from the firm's point of view, other mixed strategies. This inconvenient feature of the mixed strategy equilibrium we have is however mainly an artifact of the extreme stylization of our model and would disappear under the more realistic assumption of a continuum of possible values for 0.
- 9/ Hence for quadratic cost functions, L(v) is "very well-behaved". This suggests that the influence of higher order derivatives must be strong to induce a failure of the assumption.
- Choose  $\mu = \overline{\mu}$  and remember that  $\delta = \overline{\mu}((\overline{\vartheta} + \underline{\vartheta})/\overline{\vartheta})$  and let  $\underline{\nu}_1$  and  $((\overline{\vartheta} + \underline{\vartheta})/\overline{\vartheta})$  be small enough so that (5.14) is satisfied (and  $\delta$  is smaller than 1). We just have to satisfy condition (2) by adjusting  $\lambda$  and  $\underline{\nu}_1 < \nu$ . It is not very difficult to see that this is possible since (i)  $\forall$  1, the CP wants both firms to operate for  $\lambda$  small enough (ii)  $\forall$   $\lambda$ , the CP wants only one firm to operate for  $\underline{\nu}_1$  small enough.
- 0ne can also show that for the best scheme in the class of pooling schemes  $\Psi'(\underline{y}_1) < s$ .
- Note that the benefit of pre-production announcement exists for the old scheme and the linear scheme as well as the new scheme. The point is that in the linear and old schemes, the incentive to announce the true output is weak (in that there are other optimizing announcements) whereas truth is the only optimizing announcement in the new scheme.

- Note that these future incentive schemes have to be incentive compatible from the point of view of the planner if the firm cannot observe all the CP's new information.
- Formally bargaining sequentially over different contracts does not differ much from bargaining sequentially over a given contract.

### References

- Bonin, J. [1976], "On the Design of Managerial Incentive Structures in a Decentralized Planning Environment," American Economic Review, pp. 682-687.
- Ekern, S. [1979], "On the Soviet Incentive Model Comment." Bell Journal of Economics, Autumn.
- Fan. L. S. [1975]. "On the Reward System," American Economic Heview, pp. 226-229.
- Fudenberg, D. and J. Tirole [1983], "Sequential bargaining Under Incomplete Information," Review of Economic Studies, forthcoming.
- Gindin, S. [1970], "A Model of the Soviet Firm." Economics of Hanning, 10, No. 3, pp. 145-157.
- Guesnerie, R. and J. Seade [1980], "Non-linear Pricing in a Finite Economy," Journal of Public Economics.
- Holmstrom, B. [1979], revised 1981, "Design of Incentive Schemes and the New Soviet Incentive Model," Kellog School PP No. 456, Northwestern University.
- Kreps, D. and R. Wilson [1982], "Sequential Equilibria," Econometrica, July.
- Mirrlees, J. [1971]. "An Exploration in the Theory of Optimum Income Taxation," Review of Economic Studies, 38, pp. 175-208.
- Seade, J. [1977], "On the Shape of Optimal Tax Schedules," Journal of Public Economics 7, pp. 203-236.
- Weitzman, M. [1976], "The New Soviet Incentive Model." Bell Journal of Economics, Spring.
- Weitzman, M. [1980], "The 'Ratchet Principle' and Performance Incentives" Bell Journal of Economics, Spring.
- Yao, D. [1982], "Strategic Behavior in a Technology Forcing Regulations Setting: Automobile Emissions Control," Mimeo, GSB Stanford.

## RITORIS IN THIS SERIES

| Value | Manage | Value | Val

7 . .

### RECORD IN THIS SERVE

### Reports in this Series

- 376. "Necessary and Sufficient Conditions for Single-Peakedness Along a Linearly Ordered Set of Policy Alternatives" by P.J. Coughlin and M.J. Hinich.
- 377. "The Role of Reputation in a Regeated Agency Problem Involving Information Transmission" by W. P. Rogerson.
- 378. "Unemployment Equilibrium with Stochastic Rationing of Supplies" by Ho-mou Wu.
- 379. "Optimal Price and Income Regulation Under Uncertainty in the Model with One Producer" by M. I. Taksar.
- 380. "On the MTU Value" by Robert J. Aumann.
- 381. "Rest Invariant Estimation of a Direction Parameter with Application to Linear Functional Felationships and Factor Analysis" by T. W. Anderson, C. Stein and A. Jaman.
- 382. "Informational Equilibrium" by Robert Kast.
- 383. "Cooperative Olisopoly Equilibrium" by Mordecai Kurz.
- 384. "Reputation and Product Quality" by William P. Rogerson.
- 365. "Assiting: Perspectives from Multiperson Decision Theory" By Robert Wilson.
- 386. "Capacity Pricing" by Oren, Smith and Wilson.
- 387. "Ansequentialism and Rationality in Dynamic Choice Under Uncertainty" by P.J. Hammong.
- 388. "The Structure of Wage Contracts in Repeated Agency Models" by W. F. Frierson.
- 389. T1982 Abraham Wald Memorial Lectures, Estimating Linear Statistical be attackling by T.W. Anderson.
- 490. "Aggregates, Activities and Overheads" by W.M. Gorman.
- 391. "Double Auctions" by Robert Wilson.
- 39. "Ffficiency and Fairness in the Design of Bilateral Contracts" by  $\mathcal E$ . Honkapohja.
- 393. "Diagonality of Cost Allocation Prices" by L.J. Mirman and A. Neyman
- 39%. "General Asset Markets, Private Capital Formation, and the Existence of Temporary Walrasian Equilibrium" by P.J. Hammond
- 195. "Asymptotic Normality of the Censored and Truncated Least Absolute Deviations Estimators" by J.L. Powell
- 396. "Dominance-Colvability and Cournot Stability" by Herve Moulin
- 397. "Managerial Incentives, Investment and Aggregate Implications" by B. Holmstrom and L. Weiss

### Reports in this Series

- 398 "Generalizations of the Censored and Truncated Least Absolute Deviations Estimators" by J.L. Powell.
- 399. "Behavior Under Uncertainty and its Implications for Policy" by K.J. Arrow.
- 400. "Third-Order Efficiency of the Extended Maximum Likelihood Estimators in a Simultaneous Equation System" by K. Takeuchi and K. Morimume.
- 401. "Short-Run Analysis of Fiscal Policy in a Simple Perfect Foresight Model" by E. Judd.
- 402. "Estimation of Failure Rate From A Complete Record of Failures and a Partial Record of Mon-Failures" by K. Cuzuki.
- 403. "Applications of Semi-Regenerative Theory to Computations of Stationary Distributions of Markov Chains" by W.K. Brassmann and M.I. Taksar.
- 404. "On the Optimality of Individual Behavior in First Come Last Served Queues With Preemption and Balking" by Refael Hassin.
- 405. "Entry with Exit: An Extensive Form Treatment of Predation with Financial Constraints" by J.P. Beneit.
- 406. "Search Among Queues" by A. Glazer and R. Hassin
- 407. "The Space of Polynomials in Measures is Internal" by J. Reichers and Y. Tauman.
- 408. "Planning Under Incomplete Information and the Ratchet Effect" by X. Freixas, R. Guesnerie and J. Tirole.

# END DATE FILMED

39-83

DTIC